ELSEVIER

Stochastic Processes and their Applications 95 (2001) 1–24

stochastic processes and their applications

www.elsevier.com/locate/spa

Ergodic properties of the nonlinear filter

A. Budhiraja

Department of Statistics, University of North Carolina at Chapel Hill, Chapel Hill, NC 27599, USA

Received 31 August 2000; received in revised form 14 February 2001; accepted 18 February 2001

Abstract

In a recent work (Bhatt et al., SIAM J. Control Optim. 39 (2000) 928) various Markov and ergodicity properties of the nonlinear filter, for the classical model of nonlinear filtering, were studied. It was shown that under quite general conditions, when the signal is a Feller–Markov process with values in a complete separable metric space E then the pair process (signal, filter) is also a Feller–Markov process with state space $E \times \mathscr{P}(E)$, where $\mathscr{P}(E)$ is the space of probability measures on E. Furthermore, it was shown that if the signal has a unique invariant measure then, under appropriate conditions, uniqueness of the invariant measure for the above pair process holds within a certain restricted class of invariant measures. In many asymptotic problems concerning approximate filters (Budhiraja and Kushner, SIAM J. Control Optim. 37 (1997) 1946; 38 (2000) 1874) it is desirable to have the uniqueness of the invariant measure to hold in the class of *all* invariant measures. In this paper we first show that for a rich class of filtering problems, when the signal has a unique invariant measure, the property of "asymptotic stability" for the filter holds. Using this property of asymptotic stability we then provide sufficient conditions under which the (signal, filter) pair has a unique invariant measure. We also show that, in a certain sense, the property of asymptotic stability is necessary for the uniqueness of the invariant measure.

MSC: 60 G 35; 60 J 05; 60 H 15

Keywords: Nonlinear filtering; Invariant measures; Asymptotic stability; Measure valued processes

1. Introduction

Stochastic nonlinear filtering is one of the central areas of application of stochastic calculus (Kallianpur, 1980; Kushner, 1967; Liptser and Shiryaev, 1977). The basic object of the study is a pair of stochastic processes $(X_t, Y_t)_{t\geqslant 0}$ where (X_t) is called the signal process and (Y_t) the observation process. The central problem in nonlinear filtering is the study of the measure valued process (Π_t) which is the conditional distribution of X_t given $\sigma\{Y_s : 0 \leqslant s \leqslant t\}$. This measure valued process is called the nonlinear filter. In the classical setting of nonlinear filtering, which is considered in

E-mail address: budhiraj@email.unc.edu (A. Budhiraja).

0304-4149/01/$ - see front matter
PII: S0304-4149(01)00090-4

this paper, the signal is taken to be a Markov process with values in some Polish space E and the observations are given via the relation

$$Y_t = \int_0^t h(X_s)\,\mathrm{d}s + W_t, \tag{1.1}$$

where (W_t) is a standard d-dimensional Brownian motion independent of (X_t) and h, referred to as the observation function, is a map from $E \to \mathbb{R}^d$.

The nonlinear filter is computed using three pieces of information: the initial law of the signal (denoted hereafter as γ), the transition probability function of the Markov process (X_t, Y_t) and the observation trajectory. In most practical problems one does not have access to the exact initial law or the transition probability function. Even in the ideal situation, in order to do explicit computations various approximations need to be made. Thus it is of central importance to study the sensitivity of the filter to errors in both the initial law and the transition probability function. Most of the available work in literature focuses on the short time behavior of approximate filters. The general picture that emerges from these "short time" results is that under appropriate conditions if the errors in the parameters are small then the distance (appropriately measured) between the optimal filter and the suboptimal filter built with incorrect parameters is also small, however the bound on the distance between the two filters grows exponentially in time. These bounds suggest that over a long time interval a filter built with incorrect parameters becomes useless.

In this work, in contrast to the above-mentioned results, we are interested in the long-term behavior of the nonlinear filter. In recent years asymptotic study of the nonlinear filter has generated significant interest (Kunita, 1971, 1991; Stettner, 1989, 1991; Ocone and Pardoux, 1996; Atar and Zeitouni, 1997a, b; Cérou, 1994; Budhiraja and Ocone, 1997, 1999; Le Gland and Mevel, preprint; Atar, 1998; Clark et al., 1999; Ocone, 1999a, b; Budhiraja and Kushner, 1997, 1998, 2000, 2001; Budhiraja and Ocone, 1999; Da Prato et al., 1995; Bhatt et al., 2000; Baxendale and Liptser, 1999; Del Moral and Guionnet, 1999). In Budhiraja and Kushner (1997, 2000) it was shown that various desirable long time properties for a rich class of approximate filters hold if the pair process: (signal, filter) has a unique invariant measure.

The study of invariant measures for filtering processes was initiated by Kunita (1971). In this classic paper Kunita showed, using the uniqueness of the solution of the Kushner–Stratonovich equation, that in the classical filtering model if the signal is Feller–Markov with a compact, separable Hausdorff state space E then the optimal filter is also a Feller–Markov process with state space $\mathscr{P}(E)$, where $\mathscr{P}(E)$ is the space of all probability measures on E. Furthermore, Kunita (1971) shows that if the signal in addition has a unique invariant measure μ for which (2.11) holds then the filter $\Pi(\cdot)$ has a unique invariant measure. In subsequent papers Kunita (1991) and Stettner (1989) extended the above results to the case where the state space is a locally compact Polish space. Furthermore, in Kunita (1991) it is shown that for signals with a locally compact state space the pair process: (signal, filter) has a unique invariant measure within a certain restrictive class of invariant measures. In all the above papers (Kunita, 1971, 1991; Stettner, 1989) the observation function h is assumed to be bounded. In a recent paper (Bhatt et al., 2000) the results of Kunita–Stettner were extended to the case of unbounded h and signals with state space an arbitrary Polish space. The proofs in Bhatt et al. (2000) are of independent interest since unlike the arguments in Kunita (1971, 1991), Stettner (1989) they do not rely on the uniqueness of the solution to Kushner–Stratonovich equation.

As pointed out above, the asymptotic results in Budhiraja and Kushner (1997, 2000) crucially rely on the assumption of uniqueness of invariant measure for the pair process. Thus it becomes of central importance to obtain conditions under which this uniqueness holds. The results in Kunita (1991), Bhatt et al. (2000) on the uniqueness of the invariant measure (within a restrictive class) for the pair process are inadequate for the asymptotic study of approximate filters undertaken in Budhiraja and Kushner (1997, 2000) for the reason that the results in these latter papers require the uniqueness to hold in the class of *all* invariant measures. In view of that, in this paper, we take a different approach to this uniqueness question. To the best of our knowledge the only results (excepting the stable linear case) addressing this uniqueness problem are in Stettner (1991), Budhiraja and Kushner (2001). In Stettner (1991) it was shown that if the signal is a discrete time, finite state, aperiodic, irreducible Markov chain and the observations are given via the discrete time analog of (1.1) then the pair process admits a unique invariant measure. The results in Budhiraja and Kushner (2001) showed, for a class of discrete time signals with compact state space, that if the signal process is Feller–Markov with a unique invariant measure and furthermore the filter is "asymptotically stable" then the pair process: (signal, filter) has a unique invariant measure. A modification of the argument in Budhiraja and Kushner (2001) (see Theorem 3.6) shows that the result continues to hold for the continuous time model with the state space of the signal an arbitrary Polish space.

In view of the above result it becomes important to understand when the property of "asymptotic stability" holds. One of our aims in this paper is to obtain sufficient conditions for asymptotic stability of the filter. Roughly speaking, the property of asymptotic stability says that the distance per unit time between the optimal filter and an incorrectly initialized filter converges to 0 as time approaches ∞. More precisely, one can show that for every $v \in \mathscr{P}(E)$ there exists a family of measurable maps $\{\Lambda_t(v)\}_{t\geqslant 0}$ from $\mathscr{C} \doteq C([0,\infty):\mathbb{R}^d)$ (the space of all continuous maps from $[0,\infty)$ to $\mathbb{R}^d$) to $\mathscr{P}(E)$ such that $\Pi_t^v \doteq \Lambda_t(v)(Y_\cdot(\omega))$ represents the suboptimal filter which is constructed under the erroneous assumption that the initial law of the signal is v instead of γ. For $v \in \mathscr{P}(E)$ denote by Q_v the measure induced by (Y_t) on $\mathscr{C}$ when the Markov process (X_t) has the initial law v. Let $\mu_1, \mu_2 \in \mathscr{P}(E)$. We say that the filter is (μ_1, μ_2) asymptotically stable if for all continuous and bounded ϕ

$$\frac{1}{T}\int_0^T \mathbb{E}_{Q_{\mu_1}}[\langle \Lambda_t(\mu_1), \phi\rangle - \langle \Lambda_t(\mu_2), \phi\rangle]^2 \,\mathrm{d}t \tag{1.2}$$

converges to 0 as $T \to \infty$, where $\mathbb{E}_{Q_{\mu_1}}$ denotes the expectation with respect to the measure Q_{μ_1}.

In recent years various authors have considered the problem of asymptotic stability under different hypothesis (Ocone and Pardoux, 1996; Atar and Zeitouni, 1997a, b; Cérou, 1994; Budhiraja and Ocone, 1997, 1999; Le Gland and Mevel, preprint; Budhiraja and Kushner, 1998, 2001; Atar, 1998; Clark et al., 1999; Ocone, 1999a, b; Da Prato et al., 1995; Bhatt et al., 2000; Baxendale and Liptser, 1999; Del Moral and Guionnet, 1999). The notion of asymptotic stability that we introduce in this paper is much weaker than that studied in the above-cited papers and as a result it can be verified for a broader class of filtering problems. Another reason for us to focus on this notion of asymptotic stability is that in a certain sense (to be made precise below) it is a necessary condition for the uniqueness of the invariant measure of the pair process to hold. Furthermore, as Theorem 3.6 shows, (δ_x, μ_2) asymptotic stability for all $\mu_2 \in \mathscr{P}(E)$, x a.e. $[\mu]$, suffices for the uniqueness of the invariant measure for the pair (signal, filter) to hold, where for $x \in E$, δ_x denotes the probability measure concentrated at the point x.

The study of asymptotic stability was pioneered by Ocone and Pardoux (1996). In Ocone and Pardoux (1996) it was shown that if the observation function h is bounded and the signal is Feller–Markov on a locally compact, Polish space for which assumption (A) of Section 2 holds then for appropriate $\mu_1, \mu_2 \in \mathscr{P}(E)$ and continuous bounded real valued functions $\phi(\cdot)$ on E:

$$\lim_{t \to \infty} \mathbb{E}_{Q_{\mu_1}} [\langle \Lambda_t(\mu_1), \phi \rangle - \langle \Lambda_t(\mu_2), \phi \rangle]^2 = 0. \tag{1.3}$$

Let (T_t) denote the semi-group corresponding to the Markov process (X_t). Also, for $v \in \mathscr{P}(E)$, let

$$vT_t \doteq \int_E p(x, t, \cdot) v(\mathrm{d}x).$$

The conditions that are assumed on μ_1, μ_2 in [26] are as follows.
(A1) $\mu_1 T_t$ and $\mu_2 T_t$ converge weakly to μ as $t \to \infty$, where
(A2) $Q_{\mu_1} \ll Q_{\mu_2}$.
Our first objective in this paper is to show that if condition (A1) is weakened to the assumption that the families $\{\mu_1 T_t\}_{t \geqslant 0}$, $\{\mu_2 T_t\}_{t \geqslant 0}$ are tight then the filter is (μ_1, μ_2)-asymptotically stable. This is done in Theorem 3.1. The proof of the theorem relies on Propositions 3.10 and 3.11, the proofs of which are deferred until Section 4. Next, in Theorem 3.6, we show that if the signal admits a unique invariant measure μ and the filter is (δ_x, μ_2)-asymptotically stable for all $\mu_2 \in \mathscr{P}(E)$; x a.e. $[\mu]$ then the pair process: (signal, filter) has a unique invariant measure. The proof follows via a modification of the argument in Budhiraja and Kushner (2001). Conversely, in Theorem 3.8, it is shown that if the (signal, filter) pair admits a unique invariant measure and for some $\mu_1, \mu_2 \in \mathscr{P}(E)$ the families $\{\mu_1 T_t\}_{t \geqslant 0}$, $\{\mu_2 T_t\}_{t \geqslant 0}$ and $\{\mathbb{E}_{Q_{\mu_1}}(\Lambda_t(\mu_2))\}_{t \geqslant 0}$ are tight then the filter is (μ_1, μ_2) asymptotically stable. Note that the tightness of the above families always holds when the state space of the signal is compact. Also the tightness of the third family follows from that of the second if $Q_{\mu_1} \ll Q_{\mu_2}$. It is important to observe that in Theorems 3.6 and 3.8 assumption (A) is not made.

As an immediate consequence of the above results we show, in Theorem 3.9, that if Assumptions (A), (B1) and (B2) hold then the pair process has a unique invariant measure. Condition (B1) is sometimes referred in the literature as the statement that the Markov process is bounded in probability (Skorokhod, 1989). An important class of problems where (B2) is satisfied is when the transition probability measures $p(x, t, \mathrm{d}y)$ for the signal have nowhere vanishing densities with respect to some reference measure (cf. Proposition 3.3).

2. Notation and the filtering model

Let E be a complete separable metric space and let $(\Omega, \mathscr{F}, P)$ be a probability space. Let (X_t) be a homogeneous Markov process with values in E with transition probability function $p(x,t,B)$, i.e. for $t, \tau > 0$, $x \in E$ and $B \in \mathscr{B}(E)$

$$P(X_{t+\tau} \in B | \sigma(X_u : u \leqslant \tau)) = p(X_\tau, t, B) \text{ a.s.,} \tag{2.1}$$

where for a Polish space S, $\mathscr{B}(S)$ denotes the Borel sigma field on S. Denote the distribution of X_0 by γ, i.e.

$$\gamma = P \circ (X_0)^{-1}. \tag{2.2}$$

Denote by $\mathscr{D} \doteq D([0,\infty), E)$, the Skorohod space of E valued cadlag functions on $[0,\infty)$ and let $\xi_t(\cdot)$ be the coordinate process on $\mathscr{D}$, i.e. $\xi_t(\theta) \doteq \theta(t)$ for $\theta \in \mathscr{D}$.

We will assume that (X_t) admits a cadlag version, i.e for all $(s,x) \in [0,\infty) \times E$ there exists a probability measure $P_{s,x}$ on $\mathscr{D}$ such that for $0 \leqslant s < t < \infty$, and $U \in \mathscr{B}(E)$,

$$P_{s,x}(\xi_t \in U | \sigma(\xi_u \colon u \leqslant s)) = p(\xi_s, t-s, U) \text{ a.s. } P_{s,x} \tag{2.3}$$

and

$$P_{s,x}(\xi_u = x,\ 0 \leqslant u \leqslant s) = 1. \tag{2.4}$$

For notational simplicity, $P_{0,x}$ will hereafter be denoted as P_x.

We will also assume that the Markov process is Feller, i.e. the map $x \to P_{s,x}$ is a continuous map from E to $\mathscr{P}(\mathscr{D})$, where for a Polish space S, $\mathscr{P}(S)$ denotes the space of probability measures on S.

The observation process is given as follows:

$$Y_t = \int_0^t h(X_u)\,\mathrm{d}u + W_t, \tag{2.5}$$

where $h : E \to \mathbb{R}^d$ is a continuous and bounded mapping and (W_t) is a $\mathbb{R}^d$-valued standard Wiener process, assumed to be independent of (X_t). Denote by Π_t the conditional distribution of X_t given past and current observations, i.e. for $A \in \mathscr{B}(E)$,

$$\Pi_t(A) \doteq P(X_t \in A | \sigma\{Y_u \colon 0 \leqslant u \leqslant t\}). \tag{2.6}$$

In order to study an incorrectly initialized filter we will introduce the following canonical setting. Let (β_t) be the canonical process on $\mathscr{C} \doteq C([0,\infty) : \mathbb{R}^d)$ (the space of continuous functions from $[0,\infty)$ to $\mathbb{R}^d$), i.e. $\beta_t(\eta) \doteq \eta(t)$ for $\eta \in \mathscr{C}$. Let Q be the standard Wiener measure on $(\mathscr{C}, \mathscr{B}(\mathscr{C}))$. Also set

$$(\hat{\Omega}, \hat{\mathscr{F}}) \doteq (\mathscr{D}, \mathscr{B}(\mathscr{D})) \otimes (\mathscr{C}, \mathscr{B}(\mathscr{C}))$$

and define for $v \in \mathscr{P}(E)$, $s > 0$

$$R_{s,v} \doteq P_{s,v} \otimes Q,$$

where $P_{s,v} \in \mathscr{P}(\mathscr{D})$ is defined as

$$P_{s,v}(B) \doteq \int_E P_{s,x}(B) v(\mathrm{d}x), \quad B \in \mathscr{B}(\mathscr{D}).$$

Let $Z_t : \hat{\Omega} \to \mathbb{R}$ be the stochastic process such that for all $0 \leqslant s \leqslant t$:

$$Z_t - Z_s = \int_s^t \langle h(\xi_u), \mathrm{d}\beta_u \rangle, \text{ a.s. } R_{s,v}$$

for all $v \in \mathscr{P}(E)$, where $\langle \cdot, \cdot \rangle$ denotes the inner product in $\mathbb{R}^d$. For the existence of such a common version see Theorem 3 in Karandikar (1995). Next, for $0 \leqslant s \leqslant t$, let

$$q_{st} \doteq \exp(Z_t - Z_s - \frac{1}{2} \int_s^t \|h(\xi_u)\|^2 \, \mathrm{d}u).$$

For a Polish space S let $\mathscr{M}(S)$ denote the space of positive, finite measures on S and $\mathrm{BM}(S)$ denote the space of bounded measurable functions on S. For $f \in \mathrm{BM}(S)$ and $m \in \mathscr{M}(S)$ we will denote $\int_S f(x) \, \mathrm{d}m(x)$ by $\langle m, f \rangle$ or $m(f)$.

For $v \in \mathscr{M}(E)$ and $0 \leqslant s \leqslant t < \infty$, define a $\mathscr{M}(E)$ valued process $\Gamma_{st}(v)$ on $\mathscr{C}$ as

$$\langle \Gamma_{st}(v)(\eta), f \rangle \doteq \int_E \int_{\mathscr{D}} f(\xi_t(\theta)) q_{\mathrm{st}}(\theta, \eta) \, dP_{s,x}(\theta) \, \mathrm{d}v(x); \quad \eta - \text{a.s. } [Q]. \tag{2.7}$$

The measurability of the map $(s, t, \eta, v) \to \Gamma_{st}(v)(\eta)$ is a consequence of Theorem 3 in Karandikar (1995) which gives the measurability of the map $(t, \omega) \to Z_t(\omega)$.

Finally define for $0 \leqslant s \leqslant t$ and $v \in \mathscr{M}(E)$ a $\mathscr{P}(E)$ valued random variable $\Lambda_{st}(v)$ via the normalization of $\Gamma_{st}(v)$, i.e.

$$\Lambda_{st}(v) \doteq \frac{\Gamma_{st}(v)}{\langle \Gamma_{st}(v), 1 \rangle}.$$

Also with an abuse of notation we will sometimes denote $\Gamma_{0t}(v)$ and $\Lambda_{0t}(v)$ by $\Gamma_t(v)$ and $\Lambda_t(v)$, respectively.

As a consequence of the Kallianpur–Striebel formula (see Kallianpur, 1980) it follows that for $f \in \mathrm{BM}(E)$

$$\langle \Pi_t(\omega), f \rangle = \langle \Lambda_{0t}(\gamma)(Y_{\cdot}(\omega)), f, \rangle \quad \omega - \text{a.s. } [P]. \; t \in (0, \infty). \tag{2.8}$$

By a filter initialized incorrectly at the probability measure γ_1 we mean the $\mathscr{P}(E)$ valued process, $\Pi_t^{\gamma_1}$ defined as

$$\langle \Pi_t^{\gamma_1}(\omega), f \rangle \doteq \langle \Lambda_{0t}(\gamma_1)(Y_{\cdot}(\omega)), f \rangle, \quad f \in \mathrm{BM}(E).$$

Let $\tilde{\mathscr{F}}$ be the Q-completion of $\mathscr{B}(\mathscr{C})$ and $\tilde{\mathscr{N}}$ be the class of Q-null sets in $\tilde{\mathscr{F}}$. For $0 \leqslant s \leqslant t \leqslant \infty$, let $\mathscr{A}_s^t$ be the sub σ-fields of $\tilde{\mathscr{F}}$ defined by

$$\mathscr{A}_s^t = \sigma(\sigma(\beta_u - \beta_s \colon s \leqslant u \leqslant t) \cup \tilde{\mathscr{N}}). \tag{2.9}$$

Next we introduce the probability measure on $\mathscr{C}$ under which the canonical process has the same law as the observation process. For an arbitrary $v \in \mathscr{P}(E)$ let $Q_v \in \mathscr{P}(\mathscr{C})$ be defined by

$$\frac{\mathrm{d}Q_v}{\mathrm{d}Q} = \Gamma_t(v)(E) \quad \text{on } \mathscr{A}_0^t, \quad t \in [0, \infty). \tag{2.10}$$

It is easy to see that $PoY^{-1} = Q_\gamma$.

We now define our basic notion of asymptotic stability.

Definition 2.1. Let $\mu_1,\ \mu_2 \in \mathscr{P}(E)$. We say that the filter is (μ_1, μ_2)-asymptotically stable if for all $\phi \in C_b(E)$

$$\frac{1}{T}\int_0^T [\langle \Lambda_t(\mu_1), \phi\rangle - \langle \Lambda_t(\mu_2), \phi\rangle]^2 \,\mathrm{d}t$$

converges to 0 in Q_{μ_1}-probability as $T \to \infty$.

In the next section we will obtain conditions under which, for given $\mu_1, \mu_2 \in \mathscr{P}(E)$, (μ_1, μ_2)-asymptotic stability holds. Let (T_t) denote the semigroup corresponding to the Markov process (X_t), i.e. for $f \in \mathrm{BM}(E)$,

$$(T_t f)(x) \doteq \int_{\mathscr{D}} f(\xi_t(\theta))\,\mathrm{d}P_x(\theta).$$

For a Polish space S let $C_b(S)$ denote the space of continuous and bounded functions on S. One of the basic conditions that will be assumed in many results of this paper is the following.

(A): There is a unique invariant probability measure, μ, for the semigroup (T_t). Furthermore for all $f \in C_b(E)$:

$$\limsup_{t \to \infty} \int_E |T_t f(x) - \langle \mu, f\rangle| \mu(\mathrm{d}x) = 0. \tag{2.11}$$

3. Asymptotic stability and the uniqueness of invariant measure

In this section we first state our main result, Theorem 3.1, on asymptotic stability. We then provide sufficient conditions for Assumption (2) in Theorem 3.1 to hold. Next, we show that if the filter has appropriate asymptotic stability properties then there must be a unique invariant measure for the pair: (signal, filter). As a converse to this result, we show that if the pair process has a unique invariant measure and certain tightness conditions are satisfied then the filter is asymptotically stable. From the above results we obtain, as an immediate consequence, sufficient conditions for the uniqueness of the invariant measure to hold. We then state two propositions (Propositions 3.10 and 3.11) which play key roles in the proof of Theorem 3.1. The proofs of these propositions are deferred until Section 4. Finally, in this section, we present the proof of Theorem 3.1 using these propositions.

The main theorem of this section is the following.

Theorem 3.1. *Assume that the assumption* (A) *holds. Let* $\mu_1,\ \mu_2 \in \mathscr{P}(E)$ *be such that*:

(1) *For* $i = 1, 2$ $\{\mu_i T_t; t \geqslant 0\}$ *is a tight family of probability measures on* E.
(2) *The measure* Q_{μ_1} *is absolutely continuous with respect to* Q_{μ_2} (*we write* $Q_{\mu_1} \ll Q_{\mu_2}$).
Then the filter is (μ_1, μ_2)-*asymptotically stable, i.e. for all* $\phi \in C_b(E)$

$$\frac{1}{T}\int_0^T [\langle \Lambda_t(\mu_1), \phi\rangle - \langle \Lambda_t(\mu_2), \phi\rangle]^2 \,\mathrm{d}t$$

converges to 0 *in* Q_{μ_1}-*probability as* $T \to \infty$.

Condition (2) in the above theorem is satisfied for a rich class of problems. We give in the following proposition and the corollary below sufficient conditions for (2) to hold. First, however, we present the following lemma from Bhatt et al. (2000).

Lemma 3.2 (Bhatt et al., 2000). *Fix* $0\leqslant s<t<\infty$, $v\in\mathscr{P}(E)$. *Then*

$$\Gamma_{0t}(v)(B)(\eta)=\Gamma_{st}(\Gamma_{0s}(v)(\eta))(B)(\eta),\quad \forall B\in\mathscr{B}(E),\quad \eta-a.s.\ [Q] \tag{3.1}$$

and

$$\Lambda_{0t}(v)(B)(\eta)=\Lambda_{st}(\Lambda_{0s}(v)(\eta))(B)(\eta),\quad \forall B\in\mathscr{B}(E),\quad \eta-a.s.\ [Q]. \tag{3.2}$$

Proposition 3.3. *Let* $\mu_1,\mu_2\in\mathscr{P}(E)$. *Suppose that there exists* $\varepsilon\in(0,\infty)$ *such that, a.e.* Q, $\Lambda_\varepsilon(\mu_1)\ll\Lambda_\varepsilon(\mu_2)$. *Then* $Q_{\mu_1}\ll Q_{\mu_2}$.

Proof. Let, for $i=1,2$, $\gamma_i\doteq\Lambda_\varepsilon(\mu_i)$. By assumption, $\gamma_1\ll\gamma_2$, a.e. Q. Fix $t\in(0,\infty)$ and let $A\in\mathscr{A}_0^t$. Then

$$\begin{aligned}Q_{\mu_1}(A)&=\mathbb{E}_Q[\Gamma_{0,t}(\mu_1)(E)\mathscr{I}_A]\\&=\mathbb{E}_Q[\Gamma_{0,\varepsilon}(\mu_1)(E)\Gamma_{\varepsilon,t}(\gamma_1)(E)\mathscr{I}_A],\end{aligned} \tag{3.3}$$

where the second equality above follows from Lemma 3.2. Next note that

$$\begin{aligned}\Gamma_{\varepsilon,t}(\gamma_1)(E)&=\int_E\int_{\mathscr{D}}q_{\varepsilon,t}(\theta,\eta)\,\mathrm{d}P_{\varepsilon,x}(\theta)\,\mathrm{d}\gamma_1(x)\\&=\int_E\int_{\mathscr{D}}q_{\varepsilon,t}(\theta,\eta)\,\mathrm{d}P_{\varepsilon,x}(\theta)\frac{\mathrm{d}\gamma_1(x)}{\mathrm{d}\gamma_2(x)}\,\mathrm{d}\gamma_2(x)\\&\leqslant K\Gamma_{\varepsilon,t}(\gamma_2)(E)+\int_E\Gamma_{\varepsilon,t}(\delta_x)(E)\mathscr{I}_{\{\mathrm{d}\gamma_1(x)/\mathrm{d}\gamma_2(x)>K\}}\,\mathrm{d}\gamma_1(x),\end{aligned} \tag{3.4}$$

where $\mathscr{I}$ denotes the indicator function and $K\in(0,\infty)$ is arbitrary. Also note that

$$\begin{aligned}&\mathbb{E}_Q\left[\Gamma_{0,\varepsilon}(\mu_1)(E)\int_E\Gamma_{\varepsilon,t}(\delta_x)(E)\mathscr{I}_{\{\mathrm{d}\gamma_1(x)/\mathrm{d}\gamma_2(x)>K\}}\,\mathrm{d}\gamma_1(x)\mathscr{I}_A\right]\\&\quad\leqslant\mathbb{E}_Q\left[\Gamma_{0,\varepsilon}(\mu_1)(E)\int_E\mathscr{I}_{\{\mathrm{d}\gamma_1(x)/\mathrm{d}\gamma_2(x)>K\}}\,\mathrm{d}\gamma_1(x)\right]\\&\quad\doteq C_1(K).\end{aligned}$$

Note that $C_1(K)\to0$ as $K\to\infty$.

Furthermore for $L>0$

$$\begin{aligned}&KE_Q[\Gamma_{0,\varepsilon}(\mu_1)(E)\Gamma_{\varepsilon,t}(\gamma_2)(E)\mathscr{I}_A]\\&\quad=KE_Q\left[\Gamma_{0,\varepsilon}(\mu_2)(E)\frac{\Gamma_{0,\varepsilon}(\mu_1)(E)}{\Gamma_{0,\varepsilon}(\mu_2)(E)}\Gamma_{\varepsilon,t}(\gamma_2)(E)\mathscr{I}_A\right]\\&\quad\leqslant KL\mathbb{E}_Q[\Gamma_{0,\varepsilon}(\mu_2)(E)\Gamma_{\varepsilon,t}(\gamma_2)(E)\mathscr{I}_A]\end{aligned}$$

$$+K\mathbb{E}_Q\left[\Gamma_{0,\varepsilon}(\mu_1)(E)\Gamma_{\varepsilon,t}(\gamma_2)(E)\mathscr{I}_{\left\{\frac{\Gamma_{0,\varepsilon}(\mu_1)(E)}{\Gamma_{0,\varepsilon}(\mu_2)(E)}>L\right\}}\right]$$

$$=KL\mathbb{E}_Q[\Gamma_{0,t}(\mu_2)\mathscr{I}_A]+C_2(K,L)$$

$$=KLQ_{\mu_2}(A)+C_2(K,L),$$

where

$$C_2(K,L)\doteq K\mathbb{E}_Q\left[\Gamma_{0,\varepsilon}(\mu_1)(E)\Gamma_{\varepsilon,t}(\gamma_2)(E)\mathscr{I}_{\left\{\frac{\Gamma_{0,\varepsilon}(\mu_1)(E)}{\Gamma_{0,\varepsilon}(\mu_2)(E)}>L\right\}}\right]$$

and $C_2(K,L)\to 0$ as $L\to\infty$ for all $K\in(0,\infty)$. Note that $C_2(K,L)$ does not depend on t since

$$\mathbb{E}_Q(\Gamma_{\varepsilon,t}(\gamma_2)(E)\,|\,\mathscr{A}_0^{\varepsilon})=1$$

and $\Gamma_{0,\varepsilon}(\mu_1)$ is $\mathscr{A}_0^{\varepsilon}$ measurable. Combining the above observations we have that for all $A\in\mathscr{A}_0^t$

$$Q_{\mu_1}(A)=KLQ_{\mu_2}(A)+C_1(K)+C_2(K,L).$$

Since $t\in(0,\infty)$ is arbitrary the above relation holds for all $A\in\mathscr{C}$. Now let $A\in\mathscr{C}$ be such that $Q_{\mu_2}(A)=0$. Then the above display yields that

$$Q_{\mu_1}(A)=C_1(K)+C_2(K,L).$$

Letting $L\to\infty$ and then $K\to\infty$ we have that $Q_{\mu_1}(A)=0$. This proves the proposition. □

Corollary 3.4. *Let $\mu_1,\mu_2\in\mathscr{P}(E)$. Suppose that there exists $\varepsilon\in(0,\infty)$ such that $\mu_1T_\varepsilon\ll\mu_2T_\varepsilon$. Then $Q_{\mu_1}\ll Q_{\mu_2}$.*

Proof. In view of the above proposition it suffices to show that $\Lambda_\varepsilon(\mu_1)\ll\Lambda_\varepsilon(\mu_2)$, a.e. Q. Let $\mathscr{N}_1\subset\mathscr{C}$ be the Q null set such that for all $\eta\notin\mathscr{N}_1$, $q_{0\varepsilon}(\theta,\eta)>0$ a.s. P_{0,μ_1} and a.s. P_{0,μ_2}. Now fix $\eta\in\mathscr{N}_1^c$ and let $B\in\mathscr{B}(E)$ be such that $\langle\Lambda_\varepsilon(\mu_2)(\eta),\mathscr{I}_B\rangle=0$. This implies that,

$$\int_E\int_{\mathscr{D}}\mathscr{I}_B(\xi_\varepsilon(\theta))q_{0,\varepsilon}(\theta,\eta)\,\mathrm{d}P_x(\theta)\,\mathrm{d}\mu_2(x)=0.$$

Since $q_{0\varepsilon}(\theta,\eta)>0$, a.s. P_{μ_2} it follows that $\mu_2T_\varepsilon(B)=0$. The absolute continuity of μ_1T_ε with respect to μ_2T_ε then yields that $\mu_1T_\varepsilon(B)=0$. But this clearly implies that

$$\int_E\int_{\mathscr{D}}\mathscr{I}_B(\xi_\varepsilon(\theta))q_{0,\varepsilon}(\theta,\eta)\,\mathrm{d}P_x(\theta)\,\mathrm{d}\mu_1(x)=0.$$

Hence $\langle\Lambda_\varepsilon(\mu_1)(\eta),\mathscr{I}_B\rangle=0$. This proves that $\Lambda_\varepsilon(\mu_1)(\eta)$ is absolutely continuous with respect to $\Lambda_\varepsilon(\mu_2)(\eta)$. This proves the lemma. □

We now introduce the measure on $(\hat{\Omega}, \hat{\mathscr{F}})$ which corresponds to the law of the process $(X_t, Y_t)_{t\geqslant 0}$. For $v \in \mathscr{P}(E)$ define

$$\mathscr{K}_s^t(v) \doteq \sigma(\sigma\{\beta_u - \beta_s:\ s\leqslant u\leqslant t\} \cup \sigma\{\xi_u:\ s\leqslant u\leqslant t\} \cup \mathscr{N}), \tag{3.5}$$

where $\mathscr{N}$ is the class of all $R_{0,v}$ null sets. Now for fixed $v \in \mathscr{P}(E)$ define $\hat{R}_{0,v}$ on $(\hat{\Omega}, \hat{\mathscr{F}})$ as follows:

$$\frac{\mathrm{d}\hat{R}_{0,v}}{\mathrm{d}R_{0,v}}(\theta,\eta) \doteq q_{0t}(\theta,\eta) \quad \text{on } \mathscr{K}_0^t(v), \quad t \geqslant 0. \tag{3.6}$$

Then it can be shown that

$$\hat{R}_{0,\gamma} = Po(X., Y.)^{-1}.$$

The following theorem is taken from Bhatt et al. (2000). Let $\mathscr{F}_t$ denote the sigma field: $\sigma\{X_s, Y_s; 0\leqslant s\leqslant t\}$.

Theorem 3.5 (Bhatt et al., 2000). *Fix $v \in \mathscr{P}(E)$. Let Π_t^v be defined as*

$$\Pi_t^v(\omega) \doteq \Lambda_t(v)(Y.(\omega)).$$

Then $((X_t, \Pi_t^v), \mathscr{F}_t)$ is a $E \times \mathscr{P}(E)$ valued Feller–Markov process on $(\Omega, \mathscr{F}, P)$ with associated semigroup $\{\mathscr{S}_t\}_{0\leqslant t<\infty}$ defined as follows. For $F \in \mathrm{BM}(E \times \mathscr{P}(E))$,

$$(\mathscr{S}_t F)(x,\lambda) \doteq \mathbb{E}_{\hat{R}_{0,x}}[F(\Lambda_t(\lambda), \xi_t)]$$

for $(x,\lambda) \in E \times \mathscr{P}(E)$.

The following theorem, the essential idea of whose proof was presented in [9], gives conditions for existence and uniqueness of $(\mathscr{S}_t)$ invariant measure.

Theorem 3.6 (Budhiraja and Kushner, 2001). *Suppose that there is a unique (T_t) invariant measure μ and further suppose that the filter is (δ_x, v) asymptotically stable for all $v \in \mathscr{P}(E)$, x a.e. $[\mu]$. Then there exists a unique $(\mathscr{S}_t)$ invariant measure.*

Proof. From Theorem 6.4 of Bhatt et al. (2000) it follows that there is at least one (S_t) invariant probability measure. We note that in Theorem 6.4 of Bhatt et al. (2000) assumption (A) is made, however that assumption is only used to assure a certain uniqueness property and the existence of the invariant measure does not rely on it. Now suppose that m_1 and m_2 are two (S_t) invariant measures. We will show that for a measure determining class $\mathscr{C}_0$ of real valued functions on $(E \times \mathscr{P}(E))$ we have that for all $F \in \mathscr{C}_0$

$$\int_{E\times\mathscr{P}(E)} F(x,\alpha) m_1(\mathrm{d}x, \mathrm{d}\alpha) = \int_{E\times\mathscr{P}(E)} F(x,\alpha) m_2(\mathrm{d}x, \mathrm{d}\alpha). \tag{3.7}$$

The class $\mathscr{C}_0$ is defined as

$$\mathscr{C}_0 \doteq \{F \in C_b(E \times \mathscr{P}(E)) \,|\, \text{there exist } k \geqslant 1,\ \phi, \phi_1, \ldots, \phi_k \in C_b(E);$$
$$H \in C_b^2(\mathscr{R}^k) \text{ such that } F(x,\alpha) = \phi(x)\, H(\langle \alpha, \phi_1\rangle, \ldots, \langle \alpha, \phi_k\rangle)\},$$

where for a metric space S, $C_b(S)$ denotes the space of bounded and continuous functions and $C_b^2(\mathbb{R}^k)$ is the space of functions on $\mathbb{R}^k$ which are continuous and bounded together with their partial derivatives up to second order. The fact that $\mathscr{C}_0$ is measure determining follows from Proposition 3.4.6 of Ethier and Kurtz (1986) on noting that the class of functions $G: \mathscr{P}(E) \to \mathbb{R}$ defined as $G(\alpha) \doteq H(\langle \alpha, \phi_1 \rangle, \ldots, \langle \alpha, \phi_k \rangle)$ with $\alpha \in \mathscr{P}(E)$, $k \geqslant 1$ and H, ϕ_i as above, separates points in the space $\mathscr{P}(E)$.

Now fix $F \in \mathscr{C}_0$ and let $\phi, \phi_1, \ldots, \phi_k$ and H be as in the definition of $\mathscr{C}_0$. Then there exists a C (depending on F) such that

$$\sup_{x \in E} |F(x, \alpha_1) - F(x, \alpha_2)| \leqslant C \sum_{i=1}^{k} |\langle \alpha_1, \phi_i \rangle - \langle \alpha_2, \phi_i \rangle|. \tag{3.8}$$

Note that since there is a unique (T_t) invariant measure μ it follows that $m_i(\mathrm{d}x \times \mathscr{P}(E)) = \mu(\mathrm{d}x)$ for $i = 1, 2$. Now let μ_1, μ_2 be regular conditional probability functions such that $m_i(\mathrm{d}x, \mathrm{d}\alpha) = \mu_i(x, \mathrm{d}\alpha)\mu(\mathrm{d}x)$; $i = 1, 2$. Using the (S_t) stationarity of m_i we have that the left side of (3.7) equals, for all $T \in (0, \infty)$ and $i = 1$:

$$\frac{1}{T} \int_0^T \left(\int_E \left[\int_{\mathscr{P}(E)} S_t F(x, \alpha) \mu_i(x, \mathrm{d}\alpha) \right] \mu(\mathrm{d}x) \right) \mathrm{d}t,$$

while the right-hand side of (3.7) equals the same expression for $i = 2$. Thus

$$\begin{aligned} &\left| \int_{E \times \mathscr{P}(E)} F(x, \alpha) m_1(\mathrm{d}x, \mathrm{d}\alpha) - \int_{E \times \mathscr{P}(E)} F(x, \alpha) m_2(\mathrm{d}x, \mathrm{d}\alpha) \right| \\ &\quad \leqslant \int_E \int_{\mathscr{P}(E)} \int_{\mathscr{P}(E)} \left(\frac{1}{T} \int_0^T |S_t F(x, \alpha_1) - S_t F(x, \alpha_2)| \mathrm{d}t \right) \\ &\qquad \times \mu_1(x, \mathrm{d}\alpha_1) \mu_2(x, \mathrm{d}\alpha_2) \mu(\mathrm{d}x). \end{aligned} \tag{3.9}$$

Next using the definition of the semigroup (S_t) we have that for $x \in E$ and $\alpha_i \in \mathscr{P}(E)$; $i = 1, 2$,

$$\begin{aligned} &\frac{1}{T} \int_0^T |S_t F(x, \alpha_1) - S_t F(x, \alpha_2)| \mathrm{d}t \\ &\quad \leqslant \frac{1}{T} \int_0^T \mathbb{E}_{\hat{R}_{0,x}} |F(\xi_t, \Lambda_t(\alpha_1)) - F(\xi_t, \Lambda_t(\alpha_2))| \mathrm{d}t \\ &\quad \leqslant C \sum_{i=1}^{k} \frac{1}{T} \int_0^T \mathbb{E}_{\hat{R}_{0,x}} |\langle \Lambda_t(\alpha_1), \phi_i \rangle - \langle \Lambda_t(\alpha_2), \phi_i \rangle| \mathrm{d}t \\ &\quad \leqslant C \sum_{j=1}^{2} \sum_{i=1}^{k} \frac{1}{T} \int_0^T \mathbb{E}_{Q_x} |\langle \Lambda_t(\alpha_j), \phi_i \rangle - \langle \Lambda_t(\delta_x), \phi_i | \mathrm{d}t. \end{aligned} \tag{3.10}$$

By the assumption on asymptotic stability we have that for $i = 1, \ldots, k$,

$$\frac{1}{T} \int_0^T \mathbb{E}_{Q_x} |\langle \Lambda_t(\alpha), \phi_i \rangle - \langle \Lambda_t(\delta_x), \phi_i | \mathrm{d}t \to 0 \tag{3.11}$$

as $T \to \infty$, for all $\alpha \in \mathscr{P}(E)$ and x a.e. $[\mu]$.

Using (3.11), (3.10) in (3.9) we have via an application of dominated convergence theorem that

$$\int_{E \times \mathscr{P}(E)} F(x, \alpha) m_1(\mathrm{d}x, \mathrm{d}\alpha) = \int_{E \times \mathscr{P}(E)} F(x, \alpha) m_2(\mathrm{d}x, \mathrm{d}\alpha)$$

for all $F \in \mathscr{C}_0$. Since $\mathscr{C}_0$ is a measure determining class we have that $m_1 = m_2$. This completes the proof of the theorem. $\square$

As a converse to the above theorem we have Theorem 3.8 below. However, we first present the following lemma. Define a semigroup of operators on $\mathrm{BM}(\mathscr{P}(E))$ as follows. For $F \in \mathrm{BM}(\mathscr{P}(E))$ and $v \in \mathscr{P}(E)$

$$\begin{aligned}(\mathscr{T}_t F)(v) &\doteq \mathbb{E}_{Q_v}(F(\Lambda_t(v)))\\ &= \mathbb{E}_Q(\Gamma_t(v)(F(\Lambda_t(v)))\\ &= \mathbb{E}_Q(\Gamma_{s,t+s}(v)(F(\Lambda_{s,t+s}(v))).\end{aligned} \tag{3.12}$$

From Theorem 5.2 of Bhatt et al. (2000) we have that $(\mathscr{T}_t)$ is a Feller semigroup and from Theorem 6.2 of Bhatt et al. (2000) we have that there is a unique $(\mathscr{T}_t)$ invariant measure: $M \in \mathscr{P}(\mathscr{P}(E))$.

Following Stettner (1989), define for $v \in \mathscr{P}(E)$, $m_t^v, M_t^v \in \mathscr{P}(\mathscr{P}(E))$ as follows. For $A \in \mathscr{B}(\mathscr{P}(E))$,

$$m_t^v(A) \doteq (\mathscr{T}_t \mathscr{I}_A)(v) = \mathbb{E}_{Q_v}(\mathscr{I}_A(\Lambda_t(v)))$$

and

$$M_t^v(A) \doteq \int_E (\mathscr{T}_t \mathscr{I}_A)(\delta_x) v(\mathrm{d}x).$$

Lemma 3.7. *Let $F \in C_b(\mathscr{P}(E))$ and let $v \in \mathscr{P}(E)$ be such that $\{vT_t\}_{t \geqslant 0}$ is tight. Then*

$$\lim_{\tau \to \infty} \frac{1}{\tau} \int_0^\tau (\mathscr{T}_t F)(v)\,\mathrm{d}t = M(F).$$

Proof. The proof is similar to Proposition 3 of Stettner (1989). Recall that $(\mathscr{T}_t F)(v) = \int_{\mathscr{P}(E)} F(\pi)\,\mathrm{d}m_t^v(\pi)$. Thus the result will follow if we show that the measures $(1/\tau) \int_0^\tau m_t^v(\cdot)$ converge weakly to M. Since $(\mathscr{T}_t)$ is a Feller semigroup and M is the unique $(\mathscr{T}_t)$ invariant measure it suffices to show that the family $\{(1/\tau) \int_0^\tau m_t^v(\cdot)\,\mathrm{d}t; \tau > 0\}$ is tight. Let $\varepsilon > 0$ be arbitrary. By assumption the family $\{(1/\tau) \int_0^\tau vT_t \mathrm{d}t; \tau > 0\}$ is tight and so we have that there exists an increasing family of compact subsets of E, denoted as $\{K_n; n \geqslant 1\}$ such that

$$\frac{1}{\tau} \int_0^\tau vT_t(K_n^c)\,\mathrm{d}t \leqslant \frac{\varepsilon}{2^{2n}}.$$

Define, $J(\varepsilon) \doteq \{m \in \mathscr{P}(E)\colon m(K_n^c) \leqslant 1/2^n; n \geqslant 1\}$. Clearly $J(\varepsilon)$ is a compact set in $\mathscr{P}(E)$. Furthermore,

$$\begin{aligned}\frac{1}{\tau} \int_0^\tau m_t^v(J(\varepsilon))\,\mathrm{d}t &= \frac{1}{\tau} \int_0^\tau Q_v[\Lambda_t(v) \in J(\varepsilon)]\,\mathrm{d}t\\ &= \frac{1}{\tau} \int_0^\tau Q_v\left[\Lambda_t(v)(K_n^c) \leqslant \frac{1}{2^n}, \forall n \geqslant 1\right] \mathrm{d}t\end{aligned}$$

$$\geqslant 1 - \frac{1}{\tau}\int_0^\tau \left(\sum_{n=1}^\infty Q_v\left[\Lambda_t(v)(K_n^c) \geqslant \frac{1}{2^n}\right]\right) \mathrm{d}t$$

$$\geqslant 1 - \frac{1}{\tau}\int_0^\tau \sum_{n=1}^\infty 2^n v T_t(K_n^c)\,\mathrm{d}t$$

$$\geqslant 1 - \sum_{n=1}^\infty \frac{\varepsilon}{2^n} = 1 - \varepsilon,$$

where the next to last inequality follows on applying Chebychev's inequality and noting that $\mathbb{E}_{Q_v}(\Lambda_t(v)(K_n^c)) = vT_t(K_n^c)$. This completes the proof of the lemma. □

Theorem 3.8. *Suppose that there is a unique* (S_t) *invariant probability measure*: m. *Let* $\mu_1, \mu_2 \in \mathscr{P}(E)$ *be such that the families* $\{\mu_1 T_t\}_{t\geqslant 0}$ *and* $\{\mathbb{E}_{Q_{\mu_1}}(\Lambda_t(\mu_2))\}_{t\geqslant 0}$ *are tight. Then the filter is* (μ_1, μ_2)*-asymptotically stable.*

Proof. Let $\phi \in C_b(E)$. We need to show that

$$\frac{1}{T}\int_0^T \mathbb{E}_{Q_{\mu_1}}|\langle \Lambda_t(\mu_1), \phi\rangle - \langle \Lambda_t(\mu_2), \phi\rangle|^2\,\mathrm{d}t \tag{3.13}$$

converges to 0 as $T \to \infty$. The expression above can be rewritten as

$$\frac{1}{T}\int_0^T \mathbb{E}_{Q_{\mu_1}}(\langle \Lambda_t(\mu_1), \phi\rangle^2)\,\mathrm{d}t + \frac{1}{T}\int_0^T \mathbb{E}_{Q_{\mu_1}}(\langle \Lambda_t(\mu_2), \phi\rangle^2)\,\mathrm{d}t$$

$$-2\frac{1}{T}\int_0^T \mathbb{E}_{Q_{\mu_1}}(\langle \Lambda_t(\mu_1), \phi\rangle\langle \Lambda_t(\mu_2), \phi\rangle)\,\mathrm{d}t. \tag{3.14}$$

Observing that $\langle \Lambda_t(\mu_1), \phi\rangle = \mathbb{E}_{\hat{R}_{0,\mu_1}}(\phi(\xi_t)\,|\,\mathscr{A}_0^t)$, a.s. $\hat{R}_{0,\mu_1}$ and that $\langle \Lambda_t(\mu_2), \phi\rangle$ is $\mathscr{A}_0^t$ measurable, we have that the third term on the right-hand side of (3.14) equals

$$-2\frac{1}{T}\int_0^T \mathbb{E}_{Q_{\mu_1}}(\phi(\xi_t)\,\langle \Lambda_t(\mu_2), \phi\rangle).$$

Now let $F_\phi, G_\phi \in C_b(E \times \mathscr{P}(E))$ be defined as follows. For $(x, \alpha) \in E \times \mathscr{P}(E)$

$$F_\phi(x, \alpha) \doteq \langle \phi, \alpha\rangle^2$$

and

$$G_\phi(x, \alpha) \doteq \phi(x)\,\langle \phi, \alpha\rangle.$$

Using this notation the expression in (3.14) can be rewritten as

$$\sum_{i=1}^2 \frac{1}{T}\int_0^T \int_E S_t F_\phi(x, \mu_i)\mu_1(\mathrm{d}x) - 2\frac{1}{T}\int_0^T \int_E S_t G_\phi(x, \mu_2)\mu_1(\mathrm{d}x). \tag{3.15}$$

Define for $T \in (0,\infty)$; $\rho_T, \eta_T \in \mathscr{P}(E \times \mathscr{P}(E))$ as follows. For $A_1 \in \mathscr{B}(E)$ and $A_2 \in \mathscr{B}(\mathscr{P}(E))$:

$$\rho_T(A_1 \times A_2) \doteq \frac{1}{T}\int_0^T \int_E (S_t \mathscr{I}_{A_1 \times A_2})(x,\mu_1)\mu_1(\mathrm{d}x)$$

$$= \frac{1}{T}\int_0^T \mathbb{E}_{Q_{\mu_1}}(\mathscr{I}_{A_1}(\xi_t)\mathscr{I}_{A_2}(\Lambda_t(\mu_1)))\,\mathrm{d}t$$

and

$$\eta_T(A_1 \times A_2) \doteq \frac{1}{T}\int_0^T \int_E (S_t \mathscr{I}_{A_1 \times A_2})(x,\mu_2)\mu_1(\mathrm{d}x)$$

$$= \frac{1}{T}\int_0^T \mathbb{E}_{Q_{\mu_1}}(\mathscr{I}_{A_1}(\xi_t)\mathscr{I}_{A_2}(\Lambda_t(\mu_2)))\,\mathrm{d}t.$$

We claim that the families $\{\rho_T\}_{T\geqslant 0}$; $\{\eta_T\}_{T\geqslant 0}$ are tight in $\mathscr{P}(E \times \mathscr{P}(E))$. To see that, note that since $\{\mu_1 T_t\}_{t\geqslant 0}$ is tight in $\mathscr{P}(E)$ by assumption, it suffices to show that the family $\{(1/T)\int_0^T \mathbb{E}_{Q_{\mu_1}}(\Lambda_t(\mu_i))\,\mathrm{d}t\}_{T\geqslant 0}$ is tight in $\mathscr{P}(\mathscr{P}(E))$ for $i=1,2$. This family for $i=2$ is tight since by assumption the family $\{\mathbb{E}_{Q_{\mu_1}}(\Lambda_t(\mu_2))\}_{t\geqslant 0}$ is tight. Also, for $i=1$ this family is same as $\{\frac{1}{T}\int_0^T \mu_1 \mathscr{T}_t\,\mathrm{d}t\}_{t\geqslant 0}$ the tightness of which follows from Lemma 3.7. This proves the claim. Furthermore, observe that the expression in (3.15) can be written in terms of ρ_T and η_T as

$$\langle F_\phi, \rho_T\rangle + \langle F_\phi, \eta_T\rangle - 2\langle G_\phi, \eta_T\rangle.$$

By the Feller property of (S_t), the tightness argued above and the uniqueness of (S_t) invariant measure it follows via a routine argument that $\langle F_\phi, \rho_T\rangle \to \langle F_\phi, m\rangle$; $\langle F_\phi, \eta_T\rangle \to \langle F_\phi, m\rangle$ and $\langle G_\phi, \eta_T\rangle \to \langle G_\phi, m\rangle$ as $T\to\infty$. Hence as $T\to\infty$ the expression in (3.13) converges to

$$2\langle F_\phi, m\rangle - 2\langle G_\phi, m\rangle.$$

In order to complete the proof of the theorem it suffices to show that $\langle F_\phi, m\rangle$ equals $\langle G_\phi, m\rangle$. Next observe that

$$\langle F_\phi, m\rangle = \lim_{T\to\infty}\langle F_\phi, \rho_T\rangle$$

$$= \lim_{T\to\infty}\frac{1}{T}\int_0^T \mathbb{E}_{Q_{\mu_1}}(\langle \Lambda_t(\mu_1),\phi\rangle^2)\,\mathrm{d}t.$$

Finally note that

$$\langle G_\phi, m\rangle = \lim_{T\to\infty}\langle G_\phi, \rho_T\rangle$$

$$= \lim_{T\to\infty}\frac{1}{T}\int_0^T \mathbb{E}_{Q_{\mu_1}}(\phi(\xi_t)\langle \Lambda_t(\mu_1),\phi\rangle)\,\mathrm{d}t$$

$$= \lim_{T\to\infty}\frac{1}{T}\int_0^T \mathbb{E}_{Q_{\mu_1}}(\langle \Lambda_t(\mu_1),\phi\rangle^2)\,\mathrm{d}t,$$

where the last equality follows on observing once more that $\langle \Lambda_t(\mu_1),\phi\rangle = \mathbb{E}_{\hat{R}_{0,\mu_1}}(\phi(\xi_t)\,|\,\mathscr{A}_0^t)$, a.s. $\hat{R}_{0,\mu_1}$. Comparing the above two sets of displays we have the result. □

As an immediate consequence of Theorems 3.1, 3.5, 3.6 we have the following result.

Theorem 3.9. *Assume that assumption* (A) *holds for the filtering model. Further suppose that*:
(B1) *For all* $v \in \mathscr{P}(E)$, $\{vT_t\}_{t \geqslant 0}$ *is a tight family in* $\mathscr{P}(E)$.
(B2) $Q_x \ll Q_v$ *for all* $v \in \mathscr{P}(E)$, *x-a.e.* $[\mu]$.
Then there exists a unique $(\mathscr{S}_t)$ *invariant measure.*

The key steps in the proofs of Theorem 3.1 are Propositions 3.10 and 3.11 below. Once these are proved, Theorem 3.1 follows via a straightforward triangle inequality.

Proposition 3.10. *Suppose that assumption* (A) *holds. Let* $\mu_1 \in \mathscr{P}(E)$ *be such that* $\{\mu_1 T_t\}_{t \geqslant 0}$ *is tight. Then for all* $\phi \in C_b(E)$

$$\lim_{\tau \to \infty} \limsup_{\tau_0 \to \infty} \mathbb{E}_{Q_{\mu_1}} \left(\frac{1}{\tau_0} \int_{\tau}^{\tau_0} [\langle \Lambda_t(\mu_1), \phi \rangle - \langle \Lambda_{t-\tau,t}(\mu_1 T_{t-\tau}), \phi \rangle]^2 \, \mathrm{d}t \right) = 0. \tag{3.16}$$

Now suppose that $\mu_2 \in \mathscr{P}(E)$ *is such that* $\{\mu_2 T_t\}_{t \geqslant 0}$ *is tight and* $Q_{\mu_1} \ll Q_{\mu_2}$. *Then for all* $\phi \in \mathscr{C}_b(E)$

$$\lim_{\tau \to \infty} \limsup_{\tau_0 \to \infty} \mathbb{E}_{Q_{\mu_1}} \left(\frac{1}{\tau_0} \int_{\tau+\varepsilon}^{\tau_0} [\langle \Lambda_t(\mu_2), \phi \rangle - \langle \Lambda_{t-\tau,t}(\mu_2 T_{t-\tau-\varepsilon}), \phi \rangle]^2 \, \mathrm{d}t \right) = 0. \tag{3.17}$$

Proposition 3.11. *Let* $\tau > 0$ *be fixed. Suppose that assumption* (A) *holds. Let* $\mu_1, \mu_2 \in \mathscr{P}(E)$ *be such that the families* $\{\mu_1 T_t\}_{t \geqslant 0}$ *and* $\{\mu_2 T_t\}_{t \geqslant 0}$ *are tight. Then for all* $\phi \in C_b(E)$,

$$\mathbb{E}_{Q_{\mu_1}} \left(\frac{1}{\tau_0} \int_{\tau}^{\tau_0} [\langle \Lambda_{t-\tau,t}(\mu_1 T_{t-\tau}), \phi \rangle - \langle \Lambda_{t-\tau,t}(\mu_2 T_{t-\tau}), \phi \rangle]^2 \, \mathrm{d}t \right) \tag{3.18}$$

converges to 0 *as* $\tau_0 \to \infty$.

Assuming the above propositions we can now present the proof of Theorem 3.1.

Proof of Theorem 3.1. Let $\mu_1, \mu_2 \in \mathscr{P}(E)$ be arbitrary and let $\phi \in C_b(E)$.
Then

$$\begin{aligned}
&\mathbb{E}_{Q_{\mu_1}} \left(\frac{1}{\tau_0} \int_{0}^{\tau_0} [\langle \Lambda_t(\mu_1), \phi \rangle - \langle \Lambda_t(\mu_1), \phi \rangle]^2 \, \mathrm{d}t \right) \\
&\quad \leqslant 3 \mathbb{E}_{Q_{\mu_1}} \left(\frac{1}{\tau_0} \int_{\tau}^{\tau_0} [\langle \Lambda_t(\mu_1), \phi \rangle - \langle \Lambda_{t-\tau,t}(\mu_1 T_{t-\tau}), \phi \rangle]^2 \, \mathrm{d}t \right) \\
&\qquad + 3 \mathbb{E}_{Q_{\mu_1}} \left(\frac{1}{\tau_0} \int_{\tau}^{\tau_0} [\langle \Lambda_t(\mu_2), \phi \rangle - \langle \Lambda_{t-\tau,t}(\mu_2 T_{t-\tau}), \phi \rangle]^2 \, \mathrm{d}t \right) \\
&\qquad + 3 \mathbb{E}_{Q_{\mu_1}} \left(\frac{1}{\tau_0} \int_{\tau}^{\tau_0} [\langle \Lambda_{t-\tau,t}(\mu_1 T_{t-\tau}), \phi \rangle - \langle \Lambda_{t-\tau,t}(\mu_2 T_{t-\tau}), \phi \rangle]^2 \, \mathrm{d}t \right) \\
&\qquad + 2 \frac{\|\phi\|_{\infty}^2 (\tau)}{\tau_0},
\end{aligned}$$

where $\|\phi\|_\infty \doteq \sup_{x\in E}|\phi(x)|$. Now the theorem follows from Propositions 3.10 and 3.11 on taking $\tau_0 \to \infty$ and then $\tau \to \infty$. $\square$

4. Proofs

In this section we will present the proofs of Propositions 3.10 and 3.11. We begin with the following lemma which is an immediate consequence of Lemma 3.2.

Lemma 4.1. *Let $\mu_1 \in \mathscr{P}(E)$ and let $\tau > 0$. Then for all $t \geqslant \tau$*

$$\mathbb{E}_Q[\Gamma_{0,t}(\mu_1)(E) \,|\, \mathscr{A}^{\tau}_{t-\tau}] = \Gamma_{t-\tau,t}(\mu_1 T_{t-\tau})(E).$$

Proof. Note that from Lemma 3.2 we have that

$$\begin{aligned}\Gamma_{0,t}(\mu_1)(E) &= \Gamma_{t-\tau,t}(\Gamma_{0,t-\tau}(\mu_1))(E)\\ &= \int_E \Gamma_{t-\tau,t}(\delta_x)(E)\Gamma_{0,t-\tau}(\mu_1)(\mathrm{d}x).\end{aligned}$$

Using this observation along with the fact that under Q the sigma fields $\mathscr{A}^{t_2}_{t_1}$ and $\mathscr{A}^{t_3}_{t_2}$ are independent for $0 \leqslant t_1 \leqslant t_2 \leqslant t_3$ we have that

$$\mathbb{E}_Q[\Gamma_{0,t}(\mu_1)(E) \,|\, \mathscr{A}^{\tau}_{t-\tau}] = \int_E \Gamma_{t-\tau,t}(\delta_x)(E)\lambda(\mu_1)(\mathrm{d}x), \tag{4.1}$$

where for $\nu \in \mathscr{P}(E)$ and $A \in \mathscr{B}(E)$,

$$\begin{aligned}\lambda(\nu)(A) &\doteq \mathbb{E}_Q[\Gamma_{0,t-\tau}(\nu)(A)]\\ &= \nu T_{t-\tau}(A).\end{aligned} \tag{4.2}$$

Thus

$$\begin{aligned}\mathbb{E}_Q[\Gamma_{0,t}(\mu_1)(E) \,|\, \mathscr{A}^{\tau}_{t-\tau}] &= \int_E \Gamma_{t-\tau,t}(\delta_x)(E)\mu_1 T_{t-\tau}(\mathrm{d}x)\\ &= \Gamma_{t-\tau,t}(\mu_1 T_{t-\tau})(E).\end{aligned}$$

This proves the lemma. $\square$

Lemma 4.2. *Let $\mu_1 \in \mathscr{P}(E)$. Also let $F \in \mathrm{BM}(\mathscr{P}(E))$ and $\tau > 0$ be fixed. Then for all $t \geqslant \tau$:*

$$\mathbb{E}_{Q_{\mu_1}}(F(\Lambda_{t-\tau,t}(\mu_1 T_{t-\tau}))) = \mathbb{E}_{Q_{\mu_1}}((\mathscr{T}_\tau F)(\mu_1 T_{t-\tau})).$$

Proof. Observe that

$$\mathbb{E}_{Q_{\mu_1}}(F(\Lambda_{t-\tau,t}(\mu_1 T_{t-\tau}))) = \mathbb{E}_Q(\Gamma_{0,t}(\mu_1)(E)F(\Lambda_{t-\tau,t}(\mu_1 T_{t-\tau}))). \tag{4.3}$$

Applying Lemma 4.1 we have that the right-hand side of (4.3) equals

$$\int_{\mathscr{C}} (\mathbb{E}_Q[\Gamma_{0,t}(\mu_1)(E) \,|\, \mathscr{A}^{\tau}_{t-\tau}]F(\Lambda_{t-\tau,t}(\mu_1 T_{t-\tau})))\,\mathrm{d}Q$$

$$= \int_{\mathscr{C}} (\Gamma_{t-\tau,t}(\mu_1 T_{t-\tau})(E) F(\Lambda_{t-\tau,t}(\mu_1 T_{t-\tau}))) \, \mathrm{d}Q$$

$$= ((\mathscr{T}_\tau F)(\mu_1 T_{t-\tau})).$$

This proves the lemma. □

Lemma 4.3. *Let $\mu_1 \in \mathscr{P}(E)$ and $\tau > 0$. Fix $t > \tau$. Then for all $\phi \in \mathrm{BM}(E)$,*

$$\langle \Lambda_{t-\tau,t}(\mu_1 T_{t-\tau}), \phi \rangle = \mathbb{E}_{\hat{R}_{0,\mu_1}}[\phi(\xi_t) \mid \mathscr{A}^t_{t-\tau}], \quad a.s.\ [\hat{R}_{0,\mu_1}].$$

Proof. An application of Bayes formula yields that

$$\mathbb{E}_{\hat{R}_{0,\mu_1}}[\phi(\xi_t) \mid \mathscr{A}^t_{t-\tau}] = \frac{\mathbb{E}_{R_{0,\mu_1}}[\phi(\xi_t) q_{0t}(\theta,\eta) \mid \mathscr{A}^t_{t-\tau}]}{\mathbb{E}_{R_{0,\mu_1}}[q_{0t}(\theta,\eta) \mid \mathscr{A}^t_{t-\tau}]}. \tag{4.4}$$

We will now show that the numerator of the expression on the right-hand side almost surely equals

$$\langle \Gamma_{t-\tau,t}(\mu_1 T_{t-\tau-\varepsilon}), \phi \rangle.$$

Observe that

$$\mathbb{E}_{R_{0,\mu_1}}[\phi(\xi_t) q_{0t} \mid \mathscr{A}^t_{t-\tau}]$$
$$= \mathbb{E}_{R_{0,\mu_1}}[\phi(\xi_t)\, q_{0,t-\tau} q_{t-\tau,t} \mid \mathscr{A}^t_{t-\tau}]$$
$$= \mathbb{E}_{R_{0,\mu_1}}[\phi(\xi_t)\, q_{t-\tau,t} \mid \mathscr{A}^t_{t-\tau}]$$
$$= \int_{\mathscr{D}} \phi(\xi_t(\theta))\, q_{t-\tau,t}(\theta,\eta) \, \mathrm{d}P_{0,\mu_1}(\theta)$$
$$= \langle \Gamma_{t-\tau,t}(\mu_1 T_{t-\tau}), \phi \rangle, \tag{4.5}$$

where the second equality follows on noting that under R_{0,μ_1}, $\mathscr{A}^t_0$ is independent of $\mathscr{K}^t_{t-\tau}(\mu_1)$ and $\mathbb{E}_{R_{0,\mu_1}}(q_{0,t-\tau} \mid \mathscr{K}^t_{t-\tau}(\mu_1)) = 1$.

Combining (4.5) with (4.4) gives the result. □

Let $C_c(\mathscr{P}(E))$ be the class of all convex functions in $C_b(\mathscr{P}(E))$.

Lemma 4.4. *Let $F \in C_c(\mathscr{P}(E))$. Suppose that assumption* (A) *holds. Let M be, as before, the unique $(\mathscr{T}_t)$ invariant measure. Let $v \in \mathscr{P}(E)$ be such that $\{vT_t\}_{t \geqslant 0}$ is tight. Then*

$$\lim_{\tau \to \infty} \limsup_{\tau_0 \to \infty} \left| \frac{1}{\tau_0} \int_\tau^{\tau_0} m_\tau^{vT_{t-\tau}}(F) \, \mathrm{d}t - M(F) \right| = 0.$$

Proof. For $\tau_0 > \tau$, denote the probability measure $(1/(\tau_0 - \tau)) \int_\tau^{\tau_0} vT_{t-\tau} \, \mathrm{d}t$ by λ_{τ,τ_0}. Then clearly,

$$m_\tau^{\lambda_{\tau,\tau_0}}(F) = \mathscr{T}_\tau F \left(\frac{1}{\tau_0 - \tau} \int_\tau^{\tau_0} vT_{t-\tau} \, \mathrm{d}t \right).$$

Since F is convex we have that $\mathscr{T}_\tau F$ is convex (for a proof see Lemma 3.2 of Kunita, 1971). Therefore,

$$\begin{aligned}
m_\tau^{\lambda_{\tau,\tau_0}}(F) &= \mathscr{T}_\tau F\left(\frac{1}{\tau_0-\tau}\int_\tau^{\tau_0} vT_{t-\tau}\,\mathrm{d}t\right)\\
&\leqslant \frac{1}{\tau_0-\tau}\int_\tau^{\tau_0} \mathscr{T}_\tau F(vT_{t-\tau})\,\mathrm{d}t\\
&\leqslant \frac{1}{\tau_0-\tau}\int_\tau^{\tau_0}\left(\int_E (\mathscr{T}_\tau F)(\delta_x)vT_{t-\tau}(\mathrm{d}x)\right)\mathrm{d}t\\
&= \int_E (\mathscr{T}_\tau F)(\delta_x)\left(\frac{1}{\tau_0-\tau}\int_\tau^{\tau_0} vT_{t-\tau}(\mathrm{d}x)\,\mathrm{d}t\right)\\
&= M_\tau^{\lambda_{\tau,\tau_0}}(F).
\end{aligned}$$

This immediately yields that

$$\begin{aligned}
\left|M(F) - \frac{1}{\tau_0-\tau}\int_\tau^{\tau_0} m_\tau^{vT_{t-\tau}}(F)\,\mathrm{d}t\right| &= \left|M(F) - \frac{1}{\tau_0-\tau}\int_\tau^{\tau_0} \mathscr{T}_\tau F(vT_{t-\tau})\,\mathrm{d}t\right|\\
&\leqslant |M(F) - m_\tau^{\lambda_{\tau,\tau_0}}(F)| + |M(F) - M_\tau^{\lambda_{\tau,\tau_0}}(F)|.
\end{aligned} \tag{4.6}$$

By our assumption, $\{vT_t; t>0\}$ is tight and so the family $\{\lambda_{\tau,\tau_0} = (1/(\tau_0-\tau))\int_\tau^{\tau_0} vT_t\,\mathrm{d}t;\ \tau_0>\tau\}$ is also tight. Using the Feller property of (T_t) and the uniqueness of (T_t) invariant measure we now have that λ_{τ,τ_0} converges weakly to μ as $\tau_0\to\infty$. Furthermore since $\mathscr{T}_t$ is Feller, we have that

$$\lim_{\tau_0\to\infty} m_\tau^{\lambda_{\tau,\tau_0}}(F) = \lim_{\tau_0\to\infty} (\mathscr{T}_\tau F)(\lambda_{\tau,\tau_0}) = (\mathscr{T}_\tau F)(\mu) = m_\tau^\mu(F)$$

and

$$\lim_{\tau_0\to\infty} M_\tau^{\lambda_{\tau,\tau_0}}(F) = \lim_{\tau_0\to\infty}\int_E (\mathscr{T}_\tau F)(\delta_x)\lambda_{\tau,\tau_0}(\mathrm{d}x) = \int_E (\mathscr{T}_\tau F)(\delta_x)\mu(\mathrm{d}x) = M_\tau^\mu(F).$$

Next from the uniqueness of $(\mathscr{T}_t)$ invariant measure it follows that (cf. Section 6, Bhatt et al., 2000) $m_\tau^\mu(F)\to M(F)$ and $M_\tau^\mu(F)\to M(F)$ as $\tau\to\infty$. Now let $\varepsilon>0$ be arbitrary, then we can choose τ_ε such that for all $\tau\geqslant\tau_\varepsilon$

$$|m_\tau^\mu(F) - M(F)| + |M_\tau^\mu(F) - M(F)| \leqslant \varepsilon. \tag{4.7}$$

From (4.6) it now follows that for all $\tau\geqslant\tau_\varepsilon$,

$$\limsup_{\tau_0\to\infty}\left|M(F) - \frac{1}{\tau_0-\tau}\int_\tau^{\tau_0} m_\tau^{vT_{t-\tau}}(F)\,\mathrm{d}t\right| \leqslant \varepsilon.$$

Since F is bounded the above statement is equivalent to

$$\limsup_{\tau_0\to\infty}\left|M(F) - \frac{1}{\tau_0}\int_\tau^{\tau_0} m_\tau^{vT_{t-\tau}}(F)\,\mathrm{d}t\right| \leqslant \varepsilon.$$

Now since $\varepsilon>0$ is arbitrary, we have the lemma. $\square$

Proof of Proposition 3.10. The expectation in (3.16) can be written as

$$\frac{1}{\tau_0}\int_{\tau}^{\tau_0} \mathbb{E}_{Q_{\mu_1}}[\langle \Lambda_t(\mu_1),\phi\rangle - \langle \Lambda_{t-\tau,t}(\mu_1 T_{t-\tau}),\phi\rangle]^2\,\mathrm{d}t. \tag{4.8}$$

The expression inside the integral above can be written as

$$\begin{aligned}
&\mathbb{E}_{Q_{\mu_1}}[\langle \Lambda_t(\mu_1),\phi\rangle]^2 + \mathbb{E}_{Q_{\mu_1}}[\langle \Lambda_{t-\tau,t}(\mu_1 T_{t-\tau}),\phi\rangle]^2\\
&\quad - 2\mathbb{E}_{Q_{\mu_1}}[\langle \Lambda_t(\mu_1),\phi\rangle\,\langle \Lambda_{t-\tau,t}(\mu_1 T_{t-\tau}),\phi\rangle]\\
&\quad\quad = \mathbb{E}_{Q_{\mu_1}}[\langle \Lambda_t(\mu_1),\phi\rangle]^2 - \mathbb{E}_{Q_{\mu_1}}[\langle \Lambda_{t-\tau,t}(\mu_1 T_{t-\tau}),\phi\rangle]^2\\
&\quad\quad = (\mathscr{T}_t F_\phi)(\mu_1) - \mathbb{E}_{Q_{\mu_1}}(\mathscr{T}_\tau F_\phi)(\mu_1 T_{t-\tau}),
\end{aligned} \tag{4.9}$$

where $\tilde{F}_\phi:\mathscr{P}(E)\to\mathbb{R}$ is defined as $\tilde{F}_\phi(v)\doteq\langle\phi,v\rangle^2$ for $v\in\mathscr{P}(E)$ and the first equality in the above array follows from Lemma 4.3 while the second is a consequence of Lemma 4.2. Thus the expression in (4.8) can be written as

$$\frac{1}{\tau_0}\int_{\tau}^{\tau_0}(\mathscr{T}_t\tilde{F}_\phi)(\mu_1)\,\mathrm{d}t - \frac{1}{\tau_0}\int_{\tau}^{\tau_0}\mathbb{E}_{Q_{\mu_1}}[(\mathscr{T}_\tau\tilde{F}_\phi)(\mu_1 T_{t-\tau})]\,\mathrm{d}t. \tag{4.10}$$

From Lemma 3.7 the first expression in (4.10) converges to $\langle M,\tilde{F}_\phi\rangle$ as $\tau_0\to\infty$. Now consider the second term, then

$$\begin{aligned}
&\limsup_{\tau\to\infty}\limsup_{\tau_0\to\infty}\left|\langle M,\tilde{F}_\phi\rangle - \frac{1}{\tau_0}\int_{\tau}^{\tau_0}\mathbb{E}_{Q_{\mu_1}}(\langle m_\tau^{\mu_1 T_{t-\tau}},\tilde{F}_\phi\rangle)\,\mathrm{d}t\right|\\
&\quad\leqslant \limsup_{\tau\to\infty}\limsup_{\tau_0\to\infty}\mathbb{E}_{Q_{\mu_1}}\left|\langle M,\tilde{F}_\phi\rangle - \frac{1}{\tau_0}\int_{\tau}^{\tau_0}\langle m_\tau^{\mu_1 T_{t-\tau}},\tilde{F}_\phi\rangle\,\mathrm{d}t\right|\\
&\quad\leqslant \mathbb{E}_{Q_{\mu_1}}\left(\limsup_{\tau\to\infty}\limsup_{\tau_0\to\infty}\left|\langle M,\tilde{F}_\phi\rangle - \frac{1}{\tau_0}\int_{\tau}^{\tau_0}\langle m_\tau^{\mu_1 T_{t-\tau}},\tilde{F}_\phi\rangle\,\mathrm{d}t\right|\right)\\
&\quad = 0,
\end{aligned}$$

where the next to last step follows from the boundedness of $\tilde{F}_\phi$ and the last step is a consequence of Lemma 4.4 on observing that $-\tilde{F}_\phi$ is in $C_c(\mathscr{P}(E))$.

This shows that the second term in (4.10) also converges to $\langle M,\tilde{F}_\phi\rangle$ as $\tau_0\to\infty$ and then $\tau\to\infty$. Thus we have proved that (3.16) holds.

We now consider (3.17). The proof follows as in Ocone and Pardoux (1996). Let us define

$$G_\phi(\tau,\tau_0)\doteq\frac{1}{\tau_0}\int_0^{\tau_0}[\langle \Lambda_t(\mu_2),\phi\rangle - \langle \Lambda_{t-\tau,t}(\mu_2 T_{t-\tau-\varepsilon}),\phi\rangle]^2\,\mathrm{d}t.$$

Then

$$\begin{aligned}
\mathbb{E}_{Q_{\mu_1}}[G_\phi(\tau,\tau_0)] &= \mathbb{E}_{Q_{\mu_2}}\left[\frac{\mathrm{d}Q_{\mu_1}}{\mathrm{d}Q_{\mu_2}}G_\phi(\tau,\tau_0)\right]\\
&= K\mathbb{E}_{Q_{\mu_2}}[G_\phi(\tau,\tau_0)] + 2\|\phi\|_\infty^2\,\mathbb{E}_{Q_{\mu_2}}\left[\frac{\mathrm{d}Q_{\mu_1}}{\mathrm{d}Q_{\mu_2}}\mathscr{I}_{\frac{\mathrm{d}Q_{\mu_1}}{\mathrm{d}Q_{\mu_2}}>K}\right].
\end{aligned} \tag{4.11}$$

From the first part of the proposition we have that

$$\lim_{\tau\to\infty}\limsup_{\tau_0\to\infty}\mathbb{E}_{Q_{\mu_2}}[G_\phi(\tau,\tau_0)] = 0. \tag{4.12}$$

Using (4.11) and (4.12) we have that

$$\lim_{\tau\to\infty}\limsup_{\tau_0\to\infty}\mathbb{E}_{Q_{\mu_1}}[G_\phi(\tau,\tau_0)]\leqslant 2\|\phi\|_\infty^2\,\mathbb{E}_{Q_{\mu_2}}\left[\frac{\mathrm{d}Q_{\mu_1}}{\mathrm{d}Q_{\mu_2}}\mathscr{I}_{\frac{\mathrm{d}Q_{\mu_1}}{\mathrm{d}Q_{\mu_2}}>K}\right].$$

The proof is now completed on taking limit as $K\to\infty$ in the above expression. □

We now proceed to the proof of Proposition 3.11. Define a Markov semigroup $(\tilde{T}_t)$ on $\mathrm{BM}(E\times E)$ as follows. For $f_1,f_2\in\mathrm{BM}(E)$ define $f_1\otimes f_2$ as follows:

$$f_1\otimes f_2(x,y)\doteq f_1(x)f_2(y);\quad (x,y)\in E\times E.$$

The semigroup $\tilde{T}_t$ is now characterized by the relation

$$(\tilde{T}_t(f_1\otimes f_2))(x,y)\doteq(T_tf_1)(x)(T_tf_2)(y),\quad (x,y)\in E\times E.$$

Lemma 4.5. *Let assumption* (A) *hold. Then* $\mu\otimes\mu$ *is the unique* $(\tilde{T}_t)$ *invariant measure.*

Proof. Clearly $\mu\otimes\mu$ is one invariant measure for $(\tilde{T}_t)$. Now let v_1 be some other invariant measure for $(\tilde{T}_t)$. Define the probability measure $\tilde{v}_1$ on $(E,\mathscr{B}(E))$ as follows. For $A\in\mathscr{B}(E)$,

$$\tilde{v}_1(A)\doteq v_1(A\times E).$$

Observe that if $f\in C_b(E)$ then defining $\tilde{f}(x,y)\doteq f(x)$ we have that

$$\begin{aligned}\int_E(T_tf)(x)\tilde{v}_1(\mathrm{d}x)&=\int_{E\times E}(\tilde{T}_t\tilde{f})(x,y)v_1(\mathrm{d}x,\mathrm{d}y)\\&=\int_{E\times E}\tilde{f}(x,y)v_1(\mathrm{d}x,\mathrm{d}y)\\&=\int_E f(x)\tilde{v}_1(\mathrm{d}x).\end{aligned}$$

Hence $\tilde{v}_1$ is (T_t) invariant. By uniqueness of the (T_t) invariant measure we have that $v_1(\cdot\times E)=\tilde{v}_1(\cdot)=\mu(\cdot)$. Similarly, $v_1(E\times\cdot)=\mu(\cdot)$. Now let $f,g\in C_b(E)$. Then using the invariant properties of the measure v_1 and μ we have

$$\begin{aligned}&|v_1(f\otimes g)-\mu(f)\mu(g)|\\&\quad=\left|\int_{E\times E}(\tilde{T}_t(f\otimes g))(x,y)v_1(\mathrm{d}x\,\mathrm{d}y)-\left(\int_{E\times E}(T_tf)(x)v_1(\mathrm{d}x\,\mathrm{d}y)\right)\mu(g)\right|\\&\quad=\left|\int_{E\times E}(T_tf)(x)((T_tg)(y)-\mu(g))v_1(\mathrm{d}x,\mathrm{d}y)\right|\\&\quad\leqslant\|f\|_\infty\int_{E\times E}|T_tg(y)-\mu(g)|v_1(\mathrm{d}x,\mathrm{d}y)\\&\quad=\|f\|_\infty\int_E|T_tg(y)-\mu(g)|\mu(\mathrm{d}y).\end{aligned}$$

Finally from (2.11) the last expression converges to 0 as $t\to\infty$. Hence $v_1=\mu\otimes\mu$. □

Lemma 4.6. *Let for* $i = 1,2$, $\phi^{(i)} \in C_b(E)$. *Fix* $t > 0$ *and define* $G : E \times E \to \mathbb{R}$ *as*

$$G(x,y) \doteq \mathbb{E}_Q(\langle \Gamma_t(\delta_x), \phi^{(1)}\rangle \langle \Gamma_t(\delta_y), \phi^{(2)}\rangle), \quad x, y \in E.$$

Then $G \in C_b(E \times E)$.

Proof. Let for $i = 1,2, x_n^{(i)}$ be a sequence in E converging to $x^{(i)}$. From Theorem 5.1 of Bhatt et al. (2000) it follows that $\langle \Gamma_t(\delta_{x_n^{(i)}}), \phi^{(i)}\rangle$ converges in (Q) probability to $\langle \Gamma_t(\delta_{x^{(i)}}), \phi^{(i)}\rangle$. Also since h is bounded

$$\sup_{x\in E} \mathbb{E}_{R_0,\delta_x}(q_{0t}^p) < \infty$$

for all $p \geqslant 1$. Hence for all $\phi \in C_b(E)$, and $p \geqslant 1$

$$\sup_{x\in E} \mathbb{E}_Q[\langle \Gamma_t(\delta_x), \phi\rangle]^p < \infty.$$

This completes the proof of the lemma. $\square$

Lemma 4.7. *Let* $\tau > 0$ *be fixed. Let* $\mu_1, \mu_2 \in \mathscr{P}(E)$ *be such that the families* $\{\mu_1 T_t\}_{t\geqslant 0}$ *and* $\{\mu_2 T_t\}_{t\geqslant 0}$ *are tight. Suppose also that assumption* (A) *holds. Then for* ϕ *in* $C_b(E)$

$$\frac{1}{\tau_0}\int_\tau^{\tau_0} \mathbb{E}_Q(|\langle \Gamma_{t-\tau,t}(\mu_1 T_{t-\tau}), \phi\rangle - \langle \Gamma_{t-\tau,t}(\mu_2 T_{t-\tau}), \phi\rangle|)\,\mathrm{d}t \tag{4.13}$$

converges to 0 *as* $\tau_0 \to \infty$.

Proof. Observe that the square of the expression in (4.13) is bounded above by

$$\frac{1}{\tau_0}\int_\tau^{\tau_0} \mathbb{E}_Q(|\langle \Gamma_{t-\tau,t}(\mu_1 T_{t-\tau}), \phi\rangle - \langle \Gamma_{t-\tau,t}(\mu_2 T_{t-\tau}), \phi\rangle|^2)\,\mathrm{d}t.$$

It thus suffices to show that for $i,j = 1,2$ the limit as $\tau_0 \to \infty$ of the expression

$$\frac{1}{\tau_0}\int_\tau^{\tau_0} \mathbb{E}_Q(\langle \Gamma_{t-\tau,t}(\mu_i T_{t-\tau}), \phi\rangle \langle \Gamma_{t-\tau,t}(\mu_j T_{t-\tau}), \phi\rangle)\,\mathrm{d}t \tag{4.14}$$

exists and is independent of i,j. Now note that the expression in (4.14) can be rewritten as

$$\int_{E\times E} \mathbb{E}_Q(\langle \Gamma_\tau(\delta_x), \phi\rangle \langle \Gamma_\tau(\delta_y), \phi\rangle) m^{ij}_{\tau,\tau_0}(\mathrm{d}x\,\mathrm{d}y), \tag{4.15}$$

where

$$m^{ij}_{\tau,\tau_0} \doteq \frac{1}{\tau_0}\int_\tau^{\tau_0} (\mu_i T_{t-\tau} \otimes \mu_j T_{t-\tau})\,\mathrm{d}t.$$

Now we show that, m^{ij}_{τ,τ_0} converges weakly to $\mu \otimes \mu$ as $\tau_0 \to \infty$.

Recall that by assumption, the family $\{\mu_i T_t, t > 0; i = 1,2\}$ is tight, therefore we have that so is $\{m^{i,j}_{\tau_0}; \tau_0 > 0\}$. Now let (τ_k) be a sequence of positive numbers increasing to infinity such that $m^{i,j}_{\tau_k}$ converges to m as $k \to \infty$. We will now show that $m = \mu \otimes \mu$. Now let $f, g \in C_b(E)$. Then for $u > 0$

$$\lim_{k\to\infty} \int_{E\times E} (\tilde{T}_u(f \otimes g))(x,y) m^{i,j}_{\tau_k}(\mathrm{d}x\,\mathrm{d}y)$$

$$
\begin{aligned}
&= \lim_{k\to\infty} \frac{1}{\tau_k}\int_{\tau}^{\tau_k} (\mu_i T_{t+u-\tau} \otimes \mu_j T_{t+u-\tau})(f \otimes g)\,\mathrm{d}t \\
&= \lim_{k\to\infty} \frac{1}{\tau_k}\int_{\tau+u}^{\tau_{k+u}} (\mu_i T_{t-\tau} \otimes \mu_j T_{t-\tau})(f \otimes g)\,\mathrm{d}t \\
&= \lim_{k\to\infty} \frac{1}{\tau_k}\int_{\tau}^{\tau_k} (\mu_i T_{t-\tau} \otimes \mu_j T_{t-\tau})(f \otimes g)\,\mathrm{d}t \\
&= \lim_{k\to\infty} \langle m^{i,j}_{\tau_k}, f \otimes g\rangle \\
&= m(f \otimes g). \qquad (4.16)
\end{aligned}
$$

Also since (T_t) is a Feller semigroup the map $(x, y) \to \tilde{T}_u(f \otimes g)(x, y)$ is in $C_b(E \times E)$ and hence the extreme left-hand side expression in (4.16) converges to

$$\int_{E\times E} (\tilde{T}_u(f \otimes g))(x, y) m(\mathrm{d}x\,\mathrm{d}y).$$

Combining this observation with (4.16) we have that m is $(\tilde{T}_t)$ invariant. Hence by Lemma 4.5 it must equal to $\mu \otimes \mu$. Thus we have shown that, almost surely, m^{ij}_{τ,τ_0} converges to $\mu \otimes \mu$ as $\tau_0 \to \infty$. Thus by Lemma 4.6 we have that

$$\int_{E\times E} \mathbb{E}_Q(\langle \Gamma_\tau(\delta_x), \phi\rangle \langle \Gamma_\tau(\delta_y), \phi\rangle) m^{ij}_{\tau,\tau_0}(\mathrm{d}x\,\mathrm{d}y)$$

converges a.s. to

$$\int_{E\times E} \mathbb{E}_Q(\langle \Gamma_\tau(\delta_x), \phi\rangle \langle \Gamma_\tau(\delta_y), \phi\rangle) \mu \otimes \mu(\mathrm{d}x\,\mathrm{d}y). \qquad (4.17)$$

This proves the lemma. □

Proof of Proposition 3.11. Observe that the expression in (3.18) can be bounded above by

$$2(\|\phi\|_\infty)\frac{1}{\tau_0}\int_{\tau}^{\tau_0} \mathbb{E}_{Q_{\mu_1}}(\Xi(\mu_1, \mu_2, t))\,\mathrm{d}t,$$

where

$$\Xi(\mu_1, \mu_2, t) \doteq |\langle \Lambda_{t-\tau,t}(\mu_1 T_{t-\tau}), \phi\rangle - \langle \Lambda_{t-\tau,t}(\mu_2 T_{t-\tau}), \phi\rangle|.$$

Now note that

$$
\begin{aligned}
\mathbb{E}_{Q_{\mu_1}}(\Xi(\mu_1, \mu_2, t)) &= \mathbb{E}_Q(\Gamma_t(\mu_1)(E)\Xi(\mu_1, \mu_2, t)) \\
&= \mathbb{E}_Q(\Gamma_{t-\tau,t}(\mu_1 T_{t-\tau})(E)\Xi(\mu_1, \mu_2, t)), \qquad (4.18)
\end{aligned}
$$

where the second equality is a consequence of Lemma 4.1. Next a triangle inequality shows that

$$
\begin{aligned}
&\Gamma_{t-\tau,t}(\mu_1 T_{t-\tau})(E)\Xi(\mu_1, \mu_2, t) \\
&\quad \leqslant |\langle \Gamma_{t-\tau,t}(\mu_1 T_{t-\tau}), \phi\rangle - \langle \Gamma_{t-\tau,t}(\mu_2 T_{t-\tau}), \phi\rangle| \\
&\qquad + \|\phi\|_\infty |\Gamma_{t-\tau,t}(\mu_1 T_{t-\tau})(E) - \Gamma_{t-\tau,t}(\mu_2 T_{t-\tau})(E)|.
\end{aligned}
$$

Using this observation along with Lemma 4.7 we have that

$$\frac{1}{\tau_0}\int_{\tau}^{\tau_0} \mathbb{E}_{Q_{\mu_1}}(\Xi(\mu_1,\mu_2,t))\,\mathrm{d}t$$

converges to 0 as $\tau_0 \to \infty$. This completes the proof of the proposition. □

Acknowledgements

We would like to thank the referee for a careful study of the manuscript.

References

Atar, R., 1998. Exponential stability for nonlinear filtering of diffusion processes in non-compact domain. Ann. Probab. 26, 1552–1574.

Atar, R., Zeitouni, O., 1997a. Exponential stability for nonlinear filtering. Ann. l Inst. H. Poincare Probab. Statistique 33, 697–725.

Atar, R., Zeitouni, O., 1997b. Lyapunov exponents for finite state nonlinear filtering. SIAM J. Control Optim. 35, 36–55.

Baxendale, P., Liptser, R., 1999. Filtering stabilization under wrong initial distribution. Preprint.

Bhatt, A., Budhiraja, A., Karandikar, R., 2000. Markov property and ergodicity of the nonlinear filter. SIAM J. Control Optim. 39, 928–949.

Budhiraja, A., Kushner, H.J., 1997. Approximation and limit results for nonlinear filters over an infinite time interval. SIAM J. Control Optim. 37, 1946–1979.

Budhiraja, A., Kushner, H.J., 1998. Robustness of nonlinear filters over the infinite time interval. SIAM J. Control Optim. 36, 1618–1637.

Budhiraja, A., Kushner, H.J., 2000. Approximation and limit results for nonlinear filters over an infinite time interval: Part II, random sampling algorithms. SIAM J. Control Opt. 38, 1874–1908.

Budhiraja, A., Kushner, H.J., 2001. Monte Carlo algorithms and asymptotic problems in nonlinear filtering. In: Stochastics in Finite/Infinite Dimensions, Trends Math., Birkhuser, Boston, pp. 59–87.

Budhiraja, A., Ocone, D., 1997. Exponential stability of discrete time filters without signal ergodicity. System Control Lett. 30, 185–193.

Budhiraja, A., Ocone, D., 1999. Exponential stability in discrete time filtering for non-ergodic signals. Stochastic Processes Appl. 82, 245–257.

Cérou, F., 1994. Long time asymptotics for some dynamical noise free nonlinear filtering problems. Rapport de Recherche 2446, INRIA, December.

Clark, J.M.C., Ocone, D.L., Coumarbatch, C., 1999. Relative entropy and error bounds for filtering of Markov processes. Math. Control Signals Systems 12(4), 346–360.

Da Prato, G., Fuhrman, M., Malliavin, P., 1995. Asymptotic ergodicity for the Zakai filtering equation. C.R. Acad. Sci. Paris Ser. I 321, 613–616.

Del Moral, P., Guionnet, A., 1999. On the stability of measure valued processes. Applications to nonlinear filtering and interacting particle systems. C.R. Acad. Sci. Paris Sér. I Math. 329(5), 429–434.

Ethier, S.N., Kurtz, T.G., 1986. Markov Processes: Characterization and Convergence. Wiley, New York.

Kallianpur, G., 1980. Stochastic Filtering Theory. Springer, New York.

Karandikar, R.L., 1995. On pathwise stochastic integration. Stochastic Processes Appl. 57, 11–18.

Kunita, H., 1971. Asymptotic behavior of the nonlinear filtering errors of Markov processes. J. Multivariate Anal. 1, 365–393.

Kunita, H., 1991. Ergodic properties of nonlinear filtering processes. In: Alexander, K.C., Watkins, J.C. (Eds.), Spatial Stochastic Processes. A Festschrift in honor of Ted Harris on his seventieth birthday. Progress in Probability 19, Birkhäuser, Boston, MA.

Kushner, H.J., 1967. Dynamical equations for optimal nonlinear filtering. J. Differential Equations 3, 179–190.

Le Gland, F., Mevel, L., Geometric ergodicity in hidden Markov models. Preprint.

Liptser, R.S., Shiryaev, A.N., 1977. Statistics of Random Processes. Springer, New York.

Ocone, D., 1999a. Asymptotic stability of Benes filters. Stochastic Anal. Appl. 17(6), 1053–1074.

Ocone, D., 1999b. Entropy inequalities and entropy dynamics in nonlinear filtering of diffusion processes. In: McEneaney, W.M., Yin, G.G., Zhang, Q. (Eds.), Stochastic Analysis, Control, Optimization and Applications. A volume in honor of W.H. Fleming. Systems & Control: Foundations & Applications. Birkhäuser, Boston, MA.

Ocone, D., Pardoux, E., 1996. Asymptotic stability of the optimal filter with respect to its initial condition. SIAM J. Control Optim. 34, 226–243.

Skorokhod, A.V., 1989. Asymptotic Methods in the Theory of Stochastic Differential Equations. American Mathematical Society, Providence, AI.

Stettner, L., 1989. On invariant measures of filtering processes. In: Helmes, K., Christopeit, N., Kohlmann, M. (Eds.), Stochastic Differential Systems, Proceedings of the 4th Bad Honnef Conference, 1988, Lecture Notes in Control and Inform Sci., pp. 279–292.

Stettner, L., 1991. Invariant measures of pair: state, approximate filtering process. Colloq. Math. LXII, pp. 347–352.